Introduction

Civil Justice at the Crossroads

David Freeman Engstrom

The COVID-19 pandemic has powerfully disrupted the American legal system. Yet, as with so many other aspects of life, the pandemic was most powerful as an *accelerant* of trends already in motion. And nowhere has this been more evident in law than in the civil justice system's uptake of new legal technologies. With "legal tech" tools of all shapes and sizes gaining traction, the system, long a bastion of stasis and tradition, has begun a profound transformation.

A few examples capture the quickening pace of technological change. Start with online proceedings. Mostly a trickle pre-pandemic, the online migration of court proceedings is now a flood, to the tune of *millions* of hours each month. Judges, court administrators, lawyers, and litigants have paid switching costs in full, meaning much of the "Zooming" of litigation is likely to stick.[1] Meanwhile, after years of futurist proclamations, online dispute resolution (ODR) platforms – not live legal proceedings, but online spaces where disputants can gather, often asynchronously, and bargain their way to settlement – have now reached critical mass following a spate of COVID-era adoptions.[2] Once solely the province of e-commerce platforms like eBay, ODR is now implemented or piloted in some 200 state and local court jurisdictions, with more adoptions in the works.[3]

[1] Scott Dodson, Hon. Lee H. Rosenthal & Christopher L. Dodson, *The Zooming of Federal Civil Litigation*, 104 Judicature 13 (2020).

[2] For an accounting of ODR's advance, see David Freeman Engstrom & R. J. Vogt, *The New Judicial Governance: Courts, Data, and the Future of Civil Justice*, 72 DePaul L. Rev. (forthcoming 2023). For a graphical status update as of 2020, *see* ABA Ctr. for Innovation, Online Dispute Resolution in the United States: Data Visualizations (2020), https://www.americanbar.org/content/dam/aba/administrative/center-for-innovation/odrvisualizationreport.pdf.

[3] On ODR's private sector origins, see Ethan Katsh & Orna Rabinovich-Einy, Digital Justice: Technology and the Internet of Disputes (2017).

https://doi.org/10.1017/9781009255301.001 Published online by Cambridge University Press

New technologies are also transforming how lawyers do their day-to-day work.[4] Forced out of their offices and outside their comfort zones, lawyers have expanded their embrace of a growing menu of legal tech tools. Natural language processing (NLP), the branch of machine learning that performs text analytics and so holds the most promise in a legal system that trades in words, is advancing rapidly. Its maturation is poised to usher in a new era – with the automation of more advanced legal cognitions that supplement and even supplant lawyers' work. Going forward, lawyers will increasingly rely on AI-fired tools to perform core lawyerly tasks: reviewing documents and making privilege calls, performing legal research, choosing among arguments to lay before *this* judge, predicting case outcomes (and forecasting the cost of getting there), and even generating pleadings and papers.

Finally, the pandemic deepened what was already a growing recognition that American courts at all levels, but particularly its lower precincts, are in the grips of an access-to-justice crisis.[5] Sitting well out of view of the big-ticket, headline-grabbing litigations is a grim and less visible reality: In roughly three-quarters of civil cases filed in state courts, at least one side lacks a lawyer.[6] That means *millions* of individuals embroiled in debt-collection, eviction, and family violence and dissolution matters who cannot find lawyers and so must go it alone, often against lawyered-up adversaries. And these are only the visible litigants we can see – the tip of a very large iceberg. Below them lie *tens of millions* more who are locked out of justice entirely and choose to "lump it," or aren't aware they have a legal problem, or a legal remedy to that problem, in the first place. A growing set of states is surveying this sorry landscape and actively considering deregulating the legal services industry and relaxing the lawyers' professional monopoly in order to welcome new providers, including the software sort, into the system.[7] Online legal help, Q/A (question-answer) systems that answer legal questions in response to plain-language queries and document assembly tools that help self-represented litigants construct and file legal documents can ease pervasive access concerns – and are a final and promising legal tech frontier.

[4] For a broad overview of the legal tech industry and the academic literature that has begun to grow up around it, *see generally* David Freeman Engstrom & Jonah B. Gelbach, *Legal Tech, Civil Procedure, and the Future of Adversarialism*, 169 U. Pa. L. Rev. 1001 (2021).

[5] For a small slice of the rapid growth of concern around access to justice issues, see generally Am. Acad. of Arts & Scis., Civil Justice for All: Making Justice Accessible: Designing Legal Services for the 21st Century (2020), https://www.ncsc.org/__data/assets/pdf_file/0020/13376/civiljusticereport-2015.pdf; Access to Justice, *Daedalus* (Winter 2019).

[6] *See* Paula Hannaford-Agor, Scott Graves & Shelley Spacek Miller, Nat'l Ctr. for State Cts.: The Landscape of Civil Litigation in State Courts iv, vi, 32 (2015), https://www.ncsc.org/__data/assets/pdf_file/0020/13376/civiljusticereport-2015.pdf; Legal Servs. Corp., The Justice Gap. Measuring the Civil Legal Needs of Low-Income Americans (2017), https://www.lsc.gov/sites/default/files/images/TheJusticeGap-FullReport.pdf .

[7] *See* Engstrom & Vogt, *The New Judicial Governance*.

https://doi.org/10.1017/9781009255301.001 Published online by Cambridge University Press

Each of these pandemic-fueled tech trends will transform the civil justice system in the coming years. Debate about each is urgently needed. But to this point, much of the debate has tended toward gauzy and futurist meditations about a legal system populated by "robo-judges" and "robo-lawyers,"[8] or even an eventual state of "legal singularity," when machines perfectly predict the outcome of every case before it is filed.[9]

This volume is an antidote. It seeks to enrich – and, in places, reorient – budding debate about legal tech's implications for the civil justice system by zeroing in on the near- to medium-term, not out at a distant, hazy horizon. It does so by offering a grounded, concrete, empirically minded discussion of the *current* state of legal tech and what, *actually*, lies ahead.

Four broad questions structure the chapters that follow, as roughly allocated to four Parts. First, what are the current capacities of legal tech and where can it plausibly go over the near- to medium-term given a gauntlet of regulatory, organizational, and technological constraints? Second, what effect will legal tech's continued advance – from remote proceedings and e-discovery to legal analytics and case-outcome prediction engines – have on core features of our civil litigation system, and how might the procedural rules that structure the system adapt? Third, how can new digital technologies expand, rather than contract, access to justice for low- and moderate-income individuals who often cannot retain counsel or lack the resources or know-how to engage formal legal institutions? Fourth and finally, what aspects of judicial administration – particularly data infrastructure and accessibility – need to change in order to promote fair and responsible development of legal technologies and open the doors of justice wider for all?

Each of these questions is an essential line of inquiry if we are to seize the opportunities, and avoid the perils, of a fast-digitizing civil justice system. This volume aims to advance our understanding and lay a further research agenda as to each one.

I.1 LEGAL TECH AND THE INNOVATION ECOSYSTEM

Rigorous realism about legal tech's trajectory over the near- to medium-term must be a core part of any clear-eyed analysis of the future of the civil justice system. Part I anchors the rest of the volume by describing the current landscape of legal tech and identifying the forces that will shape its evolution from here.

[8] *See* Eugene Volokh, *Chief Justice Robots*, 68 Duke L.J. 1135, 1137–42 (2019); Milan Markovic, *Rise of the Robot Lawyers?* 61 Ariz. L. Rev. 325, 349–50 (2019).

[9] *See* Benjamin Alarie, *The Path of the Law: Towards Legal Singularity*, 66 U. Toronto L.J. 443 (2016).

https://doi.org/10.1017/9781009255301.001 Published online by Cambridge University Press

An initial task in any such effort is to survey, in understandable and accessible terms, the range of digital tools that have already begun to infuse law practice.[10] In recent decades, a growing legal-tech tool kit has steadily transformed how lawyers serve clients. An initial wave was not so different from business process optimization tools that swept into other workplaces: back-office accounting tools, document management tools, and web-based information retrieval tools like Westlaw and LexisNexis.

A second and more recent wave injected tech tools more directly into lawyers' professional workstreams. These tools focused in particular on high-volume, repetitive legal tasks formerly performed by lower-level workers – associates, contract attorneys, paralegals – within the legal services industry. E-discovery tools revolutionized the process of flagging documents for relevance and privilege that previously required manual, "eyes on" review by large work teams. But what began with keyword searches on digitized documents soon graduated to powerful machine-learning systems that, once trained on a small set of lawyer-labeled documents, can churn through millions more.[11] Another booming legal tech tool, with dozens of companies now vying for market share, identifies patterns and anomalies in large corpora of legal texts – for instance, searching thousands of contracts in order to identify where a proposed deal deviates from industry-standard provisions.[12]

A third wave, and the current state of the art, may soon eclipse these older tools through continuing advances in NLP. Potent new NLP models, such as Google's BERT or OpenAI's GPT-3, have made it possible to treat law itself as a form of data.[13] Pushing past the high-volume and repetitive tasks that were the focus of second-wave legal tech, this newest generation of legal-tech applies NLP to a wide variety of legal materials, from statutes and judicial decisions to briefs and case documents, to perform higher-order legal cognitions. The resulting applications vary. One annual effort to slice and dice the full legal tech landscape identifies a dozen different silos, from "Legal and Matter Management" to "Contract Due Diligence" and "Contract Drafting."[14] But the most promising and potentially transformative third-wave legal tech tools reduce to either of two core tasks: legal analytics and case-outcome prediction.

The most basic tools performing the first task are hunting-and-gathering tools that help lawyers corral legal materials, often accompanied by automated content

[10] *See* Engstrom & Gelbach, *Legal Tech*, at 1011–12; Engstrom & Vogt, *The New Judicial Governance*.

[11] *See* Engstrom & Gelbach, *Legal Tech*, at 1017–18; Dana A. Remus, *The Uncertain Promise of Predictive Coding*, 99 Iowa L. Rev. 1691, 1701–6 (2014).

[12] *See* Engstrom & Gelbach, *Legal Tech*, at 1017–18.

[13] *See* Michael A. Livermore & Daniel N. Rockmore, Law as Data (2018).

[14] *See 2019 Legal Tech Buyer's Guide*, LawGeex (Nov. 12, 2019), https://blog.lawgeex.com/2019-legal-tech-buyers-guide.

summaries or annotations of their relevance to the case at hand.[15] More advanced versions perform more targeted inquiries – a natural-language Q/A system for legal professionals that can return paragraphs of legal analysis at the touch of a button.[16] Still more sophisticated tools evaluate attorney work product *after* it is generated, allowing an attorney to input a draft brief and get back an analysis of questionable case law cited, apposite case law missed, or even an argument left on the table.[17] Most advanced of all, but still in development, are tools that create work product themselves, whether initial drafts of pleadings, discovery requests, or even briefs.

The second task – predicting case outcomes – is in many ways the essence of lawyering. As Justice Holmes long ago noted, law is, shorn of its pretensions, little more than a prediction as to how a court will rule.[18] Case-outcome prediction tools take various forms, but most operate by analyzing a defined set of legal and related materials – past cases, the identity of opposing counsel, the jurisdiction – and then returning a probability or set of probabilities over different outcomes, and perhaps an estimate of the time and effort needed to get there. Importantly, while outcome-prediction tools can *supplement* lawyers' work, bolstering their professional judgment with actuarial information about case prospects, they can also *supplant* lawyers. Third-wave legal tech thus holds the promise of easing access to justice concerns by arming those who cannot find or afford legal counsel with the most critical of legal information: how they would fare in court.

But just how far and how fast can these new tools go? A second vital task – and one that goes beyond merely cataloging a growing legal tech tool kit – is to map the innovation ecosystem and show how legal tech's growing capacities will shape, and be shaped by, an evolving legal services marketplace. A gauntlet of regulatory, organizational, cultural, and technological constraints will condition legal tech's continued development across different segments of the legal services landscape. Part I begins the more subtle and challenging task of reconciling these different constraints via a trio of chapter contributions by authors with leading expertise on each.

Professor Ben Barton of the University of Tennessee School of Law, a prominent expert on legal ethics and technology, maps the constraints imposed on legal tech by

[15] Michael A. Livermore et al., *Law Search in the Age of the Algorithm*, 2020 MICH. ST. L. REV. 1183 (2020).

[16] For a state-of-the-art analysis of Q/A systems, see Andrew Vold & Jack G. Conrad, *Using Transformers to Improve Answer Retrieval for Legal Questions*, 18 PROC. OF INT'L CONF. ON A.I. & L. 245 (2021).

[17] Casetext offers a pair of currently marketed examples. One is Casetext's CARA system, which ingests a whole document and outputs relevant legal authorities. *See* CASETEXT, http://www.casetext.com/cara-ai/. The other is Casetext's Compose, which purports to draft legal arguments by recommending conceptually similar legal authorities that fit a case's fact pattern as inputted by a user. *See* COMPOSE, https://compose.law.

[18] Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 458 (1897).

the tangle of rules in the United States that regulate the practice of law.[19] In "The Future of American Legal Tech: Regulation, Culture, Markets," Barton tackles a puzzle in America's market for legal technology. On one hand, our legal services industry is among the most tightly regulated in the world, suggesting infertile ground for a tech revolution. On the other is America's advanced and freewheeling market for legal tech, likely the most robust in the world. How to square the two? Barton shows how regulatory constraints – which are largely the result of an attorney discipline process overseen by state bars – is largely concentrated in the middle segment of the legal services market. It is in so-called PeopleLaw, the part of the legal services industry that serves individuals and small business, not corporate-facing BigLaw at its top or the legal aid and non-profit providers at its bottom, that regulatory constraints stymie innovation. What follows is a final and bracing insight: Given that the American legal system, for all its hidebound Burkeanism, is already the most dynamic in the world, current efforts to ease the access-to-justice crisis by loosening the rules that govern delivery of legal services, but with new modes of regulatory oversight designed to guard against consumer harm, might counter-intuitively generate more, not less, regulation than currently exists where access concerns are most acute.

Next, John Armour and Mari Sako, both from Oxford University and leading scholars of the organizational structure of the legal services industry, ask how legal technologies will proliferate in the two main "hemispheres" of legal services, BigLaw and PeopleLaw. In "Lawtech: Leveling the Playing Field in Legal Services?" they build on their past work and draw on comparative evidence from the United Kingdom to explain how the two hemispheres face very different innovation challenges.[20] For BigLaw, a principal barrier to innovation is human capital: how to attract needed talent with career-advancement opportunities and, in turn, to blend that talent in complex organizations where lawyers dominate. Another is data aggregation: how to utilize client data consistent with duties of confidentiality and conflict rules to build larger, client-spanning tools. For PeopleLaw, in stark contrast, the sticking point will be technological. PeopleLaw lawyers are largely in the business of translating law for laypeople. This "client-facing" aspect of service

[19] For his past work, *see* Benjamin H. Barton & Stephanos Bibas, Rebooting Justice: More Technology, Fewer Lawyers, and the Future of Law (2017); Benjamin H. Barton, Glass Half Full: The Decline and Rebirth of the Legal Profession (2015).

[20] For their past work, *see* John Armour & Mari Sako, *AI-Enabled Business Models in Legal Services: From Traditional Law Firms to Next-Generation Law Companies?* 7 J. Pros. & Org. 27 (2020); Richard Parnham, Mari Sako & John Armour, U. of Oxford, AI-Assisted Lawtech: Its Impact on Law Firms (2021), https://www.law.ox.ac.uk/sites/default/files/migrated/ai_final1097.pdf ; Mari Sako & Richard Parnham, U. of Oxford, Technology and Innovation in Legal Services: Final Report to the Solicitors Regulation Authority (2021), https://www.sra.org.uk/globalassets/documents/sra/research/full-report-technology-and-innovation-in-legal-services.pdf?version=4a1bfe; John Armour, Robert Parnham & Mari Sako, *Augmented Lawyering*, 2022 U. Ill. L. Rev. 71.

delivery stymies automation, because new legal tech tools, built as they are on "black box" predictive analytics, do not explain well what they do. Even as AI-fired tools automate certain repetitive and scalable tasks, a core task – explaining options and outcomes to lay clients – remains the near-exclusive competence of human lawyers.

A final contribution from Julian Nyarko of Stanford and Jens Frankenreiter of Washington University, both machine-learning experts, turns to technological constraints.[21] In "Natural Language Processing and Legal Tech," they note that new NLP techniques are already automating tasks that sit at the core of lawyering: the extraction and processing of information from a sea of unstructured text. But despite drastic improvements, they argue, NLP will, for the foreseeable future, fall well short of the "robo-judges" and "robo-lawyers" that occupy so much of the scholarly imagination. Even the most cutting-edge tools struggle with language's subtleties and law's complexities and so are not up to the legal cognitions that seasoned lawyers possess, particularly the construction of mappings of doctrine – what Nyarko and Frankenreiter call a legal "ontology." The result is that legal tech will continue to depend on humans for basic legal reasoning, reducing legal tech tools to brute-force analytic efforts once flesh-and-blood lawyers have done the important thinking.

A pair of conclusions follow from Part I's accounting of legal tech's possibilities and limits. First, legal tech's advance will not be monolithic. Rather, its incorporation into the civil justice system will be siloed, incremental, and halting – across prediction tasks, subject-matter silos, and segments of the legal services industry. In particular, it is a good bet that legal tech tools will arrive sooner, and advance most rapidly, in legal areas where data is abundant, regulated conduct takes repetitive and stereotypical forms, legal rules are inherently stable, case volumes are such that a repeat player stands to gain financially by investing, and organizational forms and culture are receptive. There is a reason why some of the most advanced legal tech tools are currently found in technocratic and self-contained areas of law (e.g., tax, labor and employment, patents) or highly routine ones (e.g., auto accidents), but not more open-ended legal contexts.

Second, Part I suggests that, over the near- to medium-term, even the most advanced legal tech tools will entail substantial lawyer engagement. Rather than full automation, legal tech may instead yield a kind of "advanced lawyering" – a spin on chess-master Gary Kasparov's notion of "advanced chess," in which human and machine ally and compete against other human-machine teams, working symbiotically, rather than merely pitting human against machine.[22] Lawyers, on this view,

[21] For their past work, see Jens Frankenreiter & Michael Livermore, *Computational Methods in Legal Analysis*, 16 Ann. Rev. L. & Soc. Sci. 39 (2020); Julian Nyarko & Sarath Sanga, A *Statistical Test for Legal Interpretation: Theory and Applications*, 38 38 J. L. Econ. & Org. 539 (forthcoming 2022).

[22] *See* Mercatus Center Podcast, *Gary Kasparov on AI, Chess, and the Future of Creativity*, Medium (May 10, 2017), https://medium.com/conversations-with-tyler/garry-kasparov-tyler-cowen-chess-iq-ai-putin-3bf28baf4dba.

may often use commoditized systems that substitute for human judgment. But a large slice of legal tech will for the foreseeable future remain customized and operate within a paradigm defined by intensive human-machine collaboration.

I.2 LEGAL TECH, LITIGATION, AND THE ADVERSARIAL SYSTEM

Part II pivots to ask: What effect will proliferating legal tech tools have on civil litigation, and what can or should be done about it? Front and center in answering these questions is what is plainly the most convulsive, and likely the most lasting, change to the legal system wrought by the COVID-19 pandemic: the migration of in-person court proceedings to remote (virtual) proceedings. However, to focus exclusively on the "Zooming" of litigation is to miss an array of legal tech innovations that likewise have the potential to reshape the litigation landscape. Indeed, the largest market segment of legal tech, fueled by the largest venture capital flows, is a growing menu of lawyer-driven tools that help lawyers do what they most commonly do: litigate cases in court on behalf of clients.[23]

The chapters in Part II recognize that the two foundational components of any litigation system are information and costs. A good starting point for envisioning the future, then, is to see how legal tech will affect each component, and also the potential trade-offs between them. Take the "Zooming" of litigation. Remote proceedings can reduce litigation costs by obviating the need for time-consuming trips to the courthouse. That's a win. Indeed, if new technologies reduce the cost of providing legal services, then those services can be brought within reach of more people who currently go without counsel. PeopleLaw might rebound and access to justice concerns might be (at least marginally) ameliorated. At the same time, the migration of in-person proceedings to online fora might also impact the *quality* of the information generated by the system. In evidentiary hearings, virtual proceedings might impact, for good and ill, crucial assessments by judges and jurors about the credibility of witnesses – and as information is impaired, judgments that rely on that information might also suffer.

That is only part of the story, for new legal technologies also hold the potential to shift the *distribution* of information and costs between litigants, thus empowering litigation's "haves" against its "have-nots," or vice versa. Legal tech's distributional effect is perhaps *the* central question in an adversarial system that, for better or worse, pits parties against one another as the best way to achieve efficient, accurate, and fair outcomes. Information, after all, is power in litigation. It confers a decisive advantage at the bargaining table and allows a litigant to make better strategic decisions, including whether to litigate in the first place. So are costs. Long-standing research in law and economics shows that a case's settlement value is, at least in part,

[23] For an introductory effort, see David Freeman Engstrom, *Digital Civil Procedure*, 169 U. Pa. L. Rev. 2177 (2021).

a function of the cost – or, rather, the projected cost – of litigating all the way to judgment.[24] New legal technologies, by increasing the quality, quantity, and cost of information for some litigants but not others, can systematically alter the litigation balance, with significant and rippling effects throughout the civil justice system.

Here too, however, predictions vary about legal tech's ultimate effect on the litigation landscape. Some see legal tech as a force-multiplier and a leveler that will allow smaller firms to do battle with larger, corporate-facing foes.[25] Legal tech can also, as with remote proceedings, reduce the cost of representation, allowing litigants to afford more of it and lawyers to tap into a latent market of individuals who currently "lump it." The darker view is that legal tech will at best replicate and at worst exacerbate existing disparities. Preliminary evidence suggests that it is corporate-facing law firms and tech companies, with privileged access to data and superior internal technical capacity, who are best positioned to develop potent legal tech tools and exploit their advantages. Bleakest of all is the possibility that legal tech will enable the "haves" to more efficiently deploy law against the "have-nots," particularly in the high-volume, low-level processes that tend to ensnare the disadvantaged and only attract low-quality counsel, if they attract counsel at all.

Five chapters size up this wider field of inquiry and plumb the ways new legal technologies will reshape the landscape of civil litigation. All are united, however, in their effort to understand how new legal technologies will affect the quantity, quality, and cost of information within the system.

In "Remote Testimonial Fact-Finding," Renee Danser and Professor James Greiner, a longtime leader of the "credibility revolution" in empirical legal studies,[26] both of Harvard and its Access to Justice Lab, anchor a team that analyzes the effect of the online migration on fact-finders. Clustering all existing studies according to the rigor of their research designs, their meta-analysis finds little evidence for the widely held concern that the online migration will adversely affect either truth-detection capacities or felt empathy toward the parties. But, not content to stop there, the authors take up the broader question of how best to study a legal system in the throes of significant change, and they offer concrete guidance on how to generate high-quality information about legal tech's impact via careful research design and a focus on causal inference.

Turning to the effects of new lawyer-driven legal tech tools, Professor Diego Zambrano of Stanford, an expert on discovery in litigation, and two

[24] For a useful overview, see Andrew F. Daughety & Jennifer F. Reinganum, *Settlement*, *in* 8 Encyclopedia of Law and Economics 386, 386–71 (Chris W. Sanchirico ed., 2nd ed. 2012).

[25] *See* Engstrom & Gelbach, *Legal Tech*, at 1037–41 (touring a growing scholarly literature).

[26] For his prior work, see D. James Greiner & Andrea Matthews, *Randomized Control Trials in the United States Legal Profession*, 12 Ann. Rev. L. & Soc. Sci. 295 (2016); D. James Greiner & Cassandra Wolos Pattanayak, *Randomized Evaluation in Legal Assistance: What Difference Does Representation (Offer and Actual Use) Make?* 121 Yale L.J. 2118 (2012); D. James Greiner et al., *The Limits of Unbundled Legal Assistance: A Randomized Study in a Massachusetts District Court and Prospects for the Future*, 126 Harv. L. Rev. 901 (2013).

co-authors – Peter Henderson and Neel Guha, both pursuing combined law and computer science degrees – zero in on e-discovery. Within civil litigation, modern discovery is the current frontline of debate about how to adapt analog procedural rules to potent machine-learning tools.[27] Their chapter, "Gamesmanship in Modern Discovery Tech," is a concrete analysis of how legal tech tools that perform core discovery tasks – often referred to collectively as "technology-assisted review" (TAR) – are revolutionizing complex litigation but also straining the system's adversarial architecture. Detractors say TAR exacerbates discovery abuse because of asymmetries in technical capacity across litigants and its inherent opacity. These critics advocate strict new protocols that compel sharing of technical inputs or outputs. Champions point to TAR's efficiencies and worry that heavy-handed regulation will stymie innovation by dissuading litigants from using TAR in the first place. Zambrano and co-authors bring much-needed light to this heated debate by offering a technical but accessible taxonomy of the ways TAR can facilitate discovery abuse. Next, they propose a menu of ways the system – especially judges, as "TAR regulators" – might plausibly police gamesmanship. The result is a powerful, discovery-specific meditation on questions that lurk in the background of all legal tech: how to balance innovation incentives against technology's potential costs, and whether new digital tools might (somewhat counterintuitively) yield a litigation system that is *more* transparent and trustworthy than the analog status quo.

In "Legal Tech and the Litigation Playing Field," Professor Nora Freeman Engstrom and Professor David Freeman Engstrom, both of Stanford, continue this line of thinking about legal tech and litigation but focus in on the distributive implications of a digitizing litigation system. Drawing on their expertise in litigation and procedure, they start with grim statistics capturing recent declines in plaintiff win rates and damages awards, which suggest that the litigation playing field has steadily slanted in favor of litigation's "haves" at the expense of its "have-nots." Next, they profile a trio of legal tech tools, beginning with TAR but then moving to a pair of frontier applications that offer a glimpse of the future: Colossus, a sophisticated case-outcome prediction tool used by the nation's largest casualty insurers to inform tens of thousands of settlements each year, and what they call the "Walmart suite," a set of tools that predict case outcomes and generate initial pleadings and papers. These innovations, they suggest, are representative of the types of tools under development by large corporate entities, tech providers, and BigLaw firms. And, if that's right, they say, the future is grim. Cutting sharply against the grain of commentary predicting legal tech will *democratize* the litigation system, they explain why such tools are, and will likely continue to be, unevenly distributed because large corporate repeat-players have privileged access to the needed technical capacity and

[27] For past relevant work by these authors, see Neel Guha, Peter Henderson & Diego A. Zambrano, *Vulnerabilities in Discovery Tech*, 35 HARV. J. L. & TECH. 581 (forthcoming 2022). Diego A. Zambrano, *Discovery as Regulation*, 120 MICH. L. REV. 71 (2020).

data, including data on case outcomes available nowhere else in a system pervaded by confidential settlements. As the American civil justice system enters the digital age, the haves will be propelled yet further ahead.

In "Litigation Outcome Prediction, Access to Justice, and Legal Endogeneity," Professor Charlotte Alexander of Georgia State, a multi-faceted legal academic and data scientist, offers a third take on legal tech's power to shape the litigation landscape via an extended analysis of the evolution and current state of the art of case-outcome prediction tools.[28] She assesses these tools' potential to fill the civil justice gap and identifies possible unintended consequences for access to justice. Her primary concern is "endogeneity" – the ability of new tech tools, particularly those that predict case outcomes, to re-shape the output of the civil justice system. In particular, she offers a concrete and gripping account of how more precise case predictions could generate hydraulic movements within the system – and, crucially, might steer risk-averse plaintiffs' counsel *away* from cases by marginalized groups and drain the system of its envelope-pushing, reformist potential. The law, already a hidebound, Burkean institution, might become more so.

In a final chapter contribution, "Toward the Participatory MDL: A Low-Tech Step to Promote Litigant Autonomy," Todd Venook of Stanford's Rhode Center on the Legal Profession and Professor Nora Freeman Engstrom turn to legal tech's impact on the big-ticket, aggregated litigations that make up a large proportion of federal court litigation. They focus on multi-district litigation (MDL), a process whereby hundreds and even thousands of individual cases are consolidated in a single federal court and litigated in bulk. Turning away from the high-tech, AI-fueled tools that occupy much of the legal tech debate, Venook and Engstrom focus instead on the potential of relatively simple court-hosted websites to promote litigant autonomy. In particular, they offer the results of their painstaking empirical study of current MDL websites. What they find is discouraging: Court websites are hard to find and frequently outdated; they lack digested, litigant-focused content; and they rarely offer litigants opportunities to attend hearings and status conferences remotely (from their home states). Finding current offerings deeply deficient, they offer principles for reform. Even a low-tech tool, Venook and Engstrom insist, can promote litigant participation and access to critical case information – but only if done right.

At least two broad conclusions emerge from this chorus of views. First, the trade-offs among core systemic values of efficiency, accuracy, and fairness as legal tech sweeps into the system will be complex and variable. More study, and more targeted study, is plainly needed if we are to understand legal tech's implications for civil litigation going forward. Second, data accessibility will be central if the legal tech

[28] For her past work on legal tech, see Charlotte S. Alexander & Mohammad Javad Feizollahi, *On Dragons, Caves, Teeth, and Claws: Legal Analytics and the Problem of Court Data Access*, *in* Computational Legal Studies: The Promise and Challenge of Data-Driven Legal Research (Ryan Whalen ed., 2019).

industry is to chart a healthy trajectory from here – sharply raising the stakes of the choices that courts are making now about data infrastructure and data governance, as taken up more fully in the remainder of the volume.

1.3 LEGAL TECH AND ACCESS TO JUSTICE

Part III of the book turns more squarely to the potential of new legal technologies to expand access to justice (A2J) for those who currently lack counsel. As previously noted, the current access crisis is deep and socially costly. A well-functioning legal infrastructure, few could deny, is critical to the achievement of collective economic and social welfare goals.[29] Distortions in the market for law and legal services can quickly undermine those goals by impairing economic activity and growth. Legal services, it seems, are far from ordinary services.[30]

As already noted, the COVID-19 pandemic, aided by a surge in consumer debt cases, brought to a boil an already-simmering debate about the high prevalence of pro se litigants within the system, placing substantial pressure on courts to innovate. Courts have moved with surprising dexterity to expand use of virtual hearings, pre-hearing diversion programs, and court-linked ODR processes. COVID-era deprivations also amplified calls to address access-to-justice concerns by deregulating the legal services industry and welcoming new providers – including the non-human, software sort – more fully into the American legal system. Finally, the pandemic turbocharged a new generation of client-facing legal tech tools designed to assist self-represented litigants by providing online legal guides, document assembly tools, and case management systems to improve access to justice for people facing problems with housing, money, family, work, and other justiciable issues.[31]

Each of these currents raises important questions about the degree to which the uptake of new legal technologies, and the rules governing their use, will widen or, to the contrary, narrow the ability of those with civil justice needs to vindicate their rights. Four chapter contributions capture the most critical aspects of legal tech and access to justice and help lay a research agenda for the future.

One crucial line of inquiry is how individuals with civil justice needs use, and could better use, the internet to access legal information and services. In "The Supply and Demand of Legal Help on the Internet," Margaret Hagan, an expert on user-centered legal design and director of Stanford's Legal Design Lab, offers both a

[29] *See* Gillian K. Hadfield, *Legal Markets*, 60 J. Econ. Lit. (forthcoming 2022).

[30] *Id.* at 19; *see also* Gillian K. Hadfield, Rules for a Flat World: Why Humans Invented Law and How to Reinvent It for a Complex Global Economy (2017); Frederick Wilmot-Smith, Equal Justice: Fair Legal Systems in an Unfair World (2019).

[31] For an overview, see Rebecca Sandefur, Am. Bar Found., Legal Tech for Non-lawyers: Report of the Survey of US Legal Technologies (2019), http://www.americanbarfoundation.org/uploads/cms/documents/report_us_digital_legal_tech_for_nonlawyers.pdf.

history lesson and a novel empirical study of the current state of online legal search.[32] The status quo, she observes, follows decades of promises that new digital technologies, anchored by the Internet, could dramatically improve access to justice. But, similar to Venook and Engstrom, she notes that insufficient attention has been paid to the actual user experience – how to make people's "legal journeys" less burdensome and more effective. Hagan then mounts a first-of-its-kind effort to gauge present-day supply of and demand for online legal help by measuring website traffic across the mix of commercial, court-linked, and public interest websites that vie for eyeballs online. Commercial sites, she finds, dominate over the more limited ecosystem of court-linked and public interest online resources, and yet commercial sites often engage in questionable practices, including the baiting of users with incomplete information and then charging for more. Search engine algorithms likely bolster that dominance. Policy implications abound: What role should search engines play to promote access to quality legal information? Could they, or should they, privilege trustworthy sources? Might there be scope for public-private partnerships, or even a regulatory role, to ensure that online searches return trustworthy and actionable legal information?

A second critical line of inquiry returns to the effect of the "Zooming" of litigation, but from an access-to-justice perspective. In "Digital Inequalities and Access to Justice: Dialing into Zoom Court Unrepresented," Professor Victor Quintanilla of Indiana University, a leading scholarly voice on civil justice design and pro se litigants, leads a research team that explores how low-income persons experience online proceedings in the everyday.[33] To date, commentators have lauded remote proceedings because of their capacity to lower litigation costs and provide conveniences. Drawing on data painstakingly compiled from Indiana courts, however, Quintanilla and co-authors paint a more complex and troubling portrait. In particular, they document persistent and pressing representational asymmetries among low-income persons that compound with digital divides in virtual court. They then identify social-psychological mechanisms through which the resulting technological asymmetries construct the status of self-represented individuals in the eyes of judges, lawyers, and even the litigants themselves. The result is a sharply different view than that offered by Danser, Greiner, and co-authors in Chapter 4, who see little basis for concern about the transition to remote testimony. Indeed, the starkness of the digital divide in evidence in Indiana gives reason to worry that the online migration might open the (virtual) courthouse doors wider only to relegate some to the (virtual) basement. The juxtaposition of these two

[32] For an example of her past work, see Margaret Hagan, *The User Experience of the Internet as a Legal Help Service: Defining Standards for the Next Generation of User-Friendly Online Legal Services*, 20 Va. J.L. & Tech. 394 (2016).

[33] For his past work on these topics, see Victor D. Quintanilla, *Human-Centered Civil Justice Design*, 121 Penn St. L. Rev. 745 (2017); Victor D. Quintanilla et al., *The Signaling Effect of Pro Se Status*, 42 Law & Soc. Inquiry 1091 (2017).

chapters, and their diverging approaches, suggests a fertile and important area for future inquiry.

Part III closes with a pair of dueling chapter contributions on the possibilities and limits of ODR – plainly ground-zero in the debate over the future of legal tech and access to justice. In "Online Dispute Resolution and the End of Adversarial Justice?" Professor Norm Spaulding of Stanford, an expert in litigation procedure and the history of the legal profession, canvasses key moments in the history of technology and shows that the downsides of innovations, and the alternatives available at their adoption, are often conveniently forgotten once path-dependence and the leverage of market dominance sets in.[34] This concern is especially acute for legal technologies designed to supplant a civil litigation system that has drawn intense criticism, only some of it earned, about excessive cost, delay, and adversarialism. Spaulding sharply critiques optimistic rhetoric around ODR and decries what he sees as undefended assumptions about case value and complexity, litigant needs, and the lawyer's professional role. In particular, he shows how ODR, at least in its current guise, is largely a technology of compliance, not adjudication, that greases the wheels of the collection efforts of companies and governments against marginalized populations. His critique counsels extreme caution in adopting, designing, and overseeing proliferating ODR platforms.

In "Using ODR Platforms to Level the Playing Field: Improving Pro Se Litigation through ODR Design," Professor J.J. Prescott of the University of Michigan, a lawyer, economist, and founder of Matterhorn, a leading ODR company, pushes back against Professor Spaulding's sharp critique.[35] It is unfair, he says, to compare court adjudication at its best to ODR at its worst. Warts and all, current ODR systems might be better than nothing in a system with pervasive access problems. But Prescott also concedes the shortcomings of ODR in its current guise and asks how we might address them. In particular, while existing ODR platforms might usefully simplify and clarify legal steps, lawyerless litigants still proceed without experience, expertise, guardrails, or the ability to gauge risk or outcomes. It is thus time, Prescott argues, to begin the work of thinking about what the next generation of ODR technology, or ODR 2.0, might look like. In particular, we might inject ODR with a dose of data science, arming self-represented litigants with the information, including legal options and predictions about their case, they might receive from

[34] For past work on AI and legal tech, see Norman Spaulding, *Is Human Judgment Necessary? Artificial Intelligence, Algorithmic Governance, and the Law*, *in* The Oxford Handbook of Ethics of AI (Markus D. Duber et al. eds., 2020).

[35] For his past work on ODR, see Maximilian A. Bulinski & J.J. Prescott, *Online Case Resolution Systems: Enhancing Access, Fairness, Accuracy, and Efficiency*, 21 Mich. J. Race & L. 205 (2016); Avital Mentovich, J.J. Prescott & Orna Rabinovich-Einy, *Are Litigation Outcome Disparities Inevitable? Courts, Technology, and the Future of Impartiality*, 71 Ala. L. Rev. 893 (2020). For Matterhorn, see Matterhorn, https://getmatterhorn.com/.

a lawyer were one available. Enhanced ODR, he notes, is unlikely to render lawyer representation obsolete, but next-generation platforms can dramatically reduce the justice gap and, on key dimensions, be a significant improvement over the status quo.

As with other Parts of this volume, these contributions cover only a slice of a fast-moving field. "Regulatory reform," described in passing by Professor Barton in Chapter 1, is gaining momentum. Two states, Utah and Arizona, have already relaxed the usual rules governing legal services. If more states follow – and several states, including California, are actively considering doing so – the result would be tectonic.[36] Another wild card is litigation invoking antitrust law and the First Amendment against the rules that sustain the lawyer's professional monopoly. Using these time-worn principles, federal courts could forcibly loosen constraints on the provision of legal services even if states will not. Any of the above could usher new providers, including software-based ones, into the system.

Whether either of these things comes to pass, the above chapter contributions combine to form a generalizable insight about legal tech and access to justice. As legal tech of all sorts proliferates, we must continually ask what innovations will increase access *to*, and also what kind of justice, and what kind of legal subject, they will combine to deliver. Only by continually asking and answering these questions can judges, court administrators, lawyers, and policy makers guard against the risk that new technologies will see rapid adoption but lead to an impoverished, gutter system of justice.

1.4 COURTS, DATA, AND CIVIL JUSTICE

A fourth and final Part aims to help lawyers, judges, policy makers, and others who populate the civil justice system think about the resource that will increasingly be the lifeblood of a fast-digitizing civil justice system: data.

Digitization means datafication, and the civil justice system is generating data like never before. At the center of this data-generating machine, of course, are the courts themselves. Courts have always been data monopolists. They sit atop mountains of data that grow larger each year, fed by a ceaseless stream of litigant-filed pleadings and papers and court-issued orders and decisions. But that only begins to describe the data that a newly digitized civil justice system is throwing off. New forms of digitization, from online courts to ODR to a pandemic-fueled shift to e-filing, mean

[36] Engstrom & Vogt, *The New Judicial Governance*; *see also* Rebecca L. Sandefur, Thomas M. Clarke & James Teufel, *Seconds to Impact? Regulatory Reform, New Kinds of Legal Services, and Increased Access to Justice*, 84 Law & Contemp. Problems 69 (2021); Elizabeth Chambliss, *Evidence-Based Lawyer Regulation*, 97 Wash. U. L. Rev. 297 (2019).

a larger store of digitized legal materials than ever before that can be mined for information and insights about the system's workings.[37]

Importantly, digitization means more than just *more* data. It also means *higher-value* data that can be put to more effective use, not just by courts, but by a growing legal tech industry. As various of the chapter contributions to this volume make clear, data, and access to it, will increasingly shape the civil justice system's core workings. A key question for the future, then, is whether the civil justice system, and, in particular, the courts that oversee it, can harness new data flows and manage data's increasing centrality in ways that promote the just, equitable, and efficient administration of justice – or, alternatively, subvert it. Choices made today will have large consequences tomorrow.

Recognizing data's central importance, a final set of chapters interrogates how courts should perform their new and increasingly central role as data dispensers and data governors. And, it also examines how courts might mount more ambitious and future-looking projects, such as the creation of data trusts as a way to achieve new transparency over the workings of the system.

Data challenges track wider governance challenges. Such is the lesson of a chapter by Chief Justice McCormack of the Michigan Supreme Court, a national voice on access to justice and court reform.[38] In "The Disruption We Needed: COVID-19, Court Technology, and Access to Justice," McCormack offers a riveting, ground-level account of the remarkable strides that the normally hide-bound courts have made during the pandemic to widen access to justice through use of technology. In Michigan and elsewhere, courts moved quickly to implement ODR and diversion programs, and they built out data capacity to better inform judicial administration. But McCormack also laments the challenges that remain. The Michigan courts, as in many states, are a non-unified judicial system, with many court records under the control of elected local court clerks, steeply raising the costs of new, statewide initiatives. And that only describes the *intra*-jurisdictional challenges around data and interoperability. As many have noted, *inter*-jurisdictional differences are also severe, yielding an innovation-choking checkerboard of 14,000 federal, state, and local court jurisdictions with data and IT infrastructures that cannot talk to one another.[39] The pandemic has spurred important changes in how courts deliver justice, but much work remains to be done.

[37] For a useful overview of pandemic-fueled court digitization, *see generally* Nat'l Ctr. For State Cts., Post-Pandemic Planning: Technology Resource Guide (2020), https://www.ncsc.org/__data/assets/pdf_file/0020/42482/Post-Pandemic-Planning.pdf.

[38] For recent academic work, see Bridget Mary McCormack, *Staying Off the Sidelines: Judges as Agents for Justice System Reform*, 131 Yale L.J. Forum 175 (2021).

[39] The leading effort to address inter-jurisdictional data challenges is the National Open Court Data Standards' (NODS) project of the National Conference of State Courts. *See National Open Court Data Standards (NODS)*, Nat'l Ctr. for State Cts., https://www.ncsc.org/services-and-experts/areas-of-expertise/court-statistics/national-open-court-data-standards-nods.

Meanwhile, it is courts that will decide, and are deciding, whether to make the mountains of data generated by the legal system available to outside actors. And, so far, courts have kept a very tight grip. In "Free PACER," Professor Jonah Gelbach of Berkeley, a lawyer, economist, and leading legal empiricist, offers a plaintive appeal for Congress and the federal courts to address the kludgy paywalls behind which much federal court data sits.[40] High fees and a clunky, twentieth-century user interface hinder potentially useful research. Worse, though PACER fee revenue is often described as high, it is actually a pittance as an economic matter, meaning that scholars, policy makers, and litigants are deprived of data – but for little gain. Importantly, Gelbach's concerns are no less relevant at the state level. Even as new technologies are thickening data flows, acute budgetary pressures, exacerbated by the pandemic, make it all too tempting for state judicial administrators to monetize their holdings and place data beyond the reach of all but the most-well-heeled law firms and tech companies.

Continuing the focus on court governance, in "Technological Challenges Facing the Judiciary," Professor Albert Yoon of the University of Toronto, a lawyer, economist, and founder of Blue J Legal, an innovative legal tech company, observes that, while the legal profession has been slow to adopt technology compared to other skilled labor markets, the courts have been even slower.[41] The resulting gaps in technological sophistication both within the practicing bar, and between Bar and Bench, is exacerbating the already-wide disparities between richer and poorer litigants. Access-to-justice concerns will only worsen, Yoon predicts, as wealthier litigants use new data sources to develop increasingly potent tech tools and then exploit the advantages they confer against a technologically backward judiciary. Moreover, the judiciary will find itself at a disadvantage relative to more technologically sophisticated parties, if for no other reason than it will lag in its ability to efficiently make decisions. The judiciary, Yoon argues, can and should reverse that trend. Needed steps include democratizing court data by making it more publicly available and increased willingness among judges to develop and use data-based tools to inform their own judicial decision-making.

Finally, in "The Civil Justice Data Gap," Professor Tanina Rostain and Amy O'Hara of Georgetown, both pioneers in thinking about data governance and the legal system, ask how we might reimagine the collection, sharing, and analysis of data to make the civil justice system more accountable to other government

[40] For his prior work using federal court data to analyze civil justice problems, see Jonah B. Gelbach, *Can We Learn Anything about Pleading Changes from Existing Data?* 44 Int'l Rev. L. & Econ. 72 (2015); Jonah B. Gelbach, *Can the Dark Arts of the Dismal Science Shed Light on the Empirical Reality of Civil Procedure?* 2 Stan. J. Complex Litig. 223 (2014).

[41] For prior relevant work, see Yoon, *Post-Modern Lawyer*. Professor Yoon is a founder of Blue J Legal, a legal tech company specializing in tax and labor employment matters. *See* Blue J. Legal, https://www.bluej.com/.

institutions, litigants, and the public at large.[42] They tackle those questions, first, by identifying vital lines of research that would benefit from better data. Among these are the ways racism shapes court processes and outcomes, the consequences of a lack of representation on litigants' experiences and case outcomes within the civil justice system, and the upstream determinants and downstream consequences of judicial entanglement. Having made the case for how better data might be put to salutary use, they report on their ongoing efforts to design and build a "civil justice data commons" that might overcome these barriers by hosting documented, harmonized data and even providing tools and analytic software to conduct research. A workable data commons would enrich policy research and yield operational insights for courts and civil justice institutions seeking to improve equity and service delivery.

Part IV's collected contributions make two points abundantly clear: First, as the civil justice system continues the process of digitization, the health of the system will increasingly turn on the health of its data ecosystem. Second, because data will be increasingly central to civil justice, courts will necessarily play a critically important but novel role as data governors. Performing that new data governance role, and navigating the opportunities presented by a newly digitized system while avoiding its many perils, will require careful regulation to ensure that data is distributed and used in fair, equitable, secure, and privacy-protecting ways.

* * *

More than a hundred years ago, Justice Holmes, in *The Path of the Law*, wrote: "For the rational study of the law the black-letter man may be the man of the present, but the man of the future is the man of statistics and the master of economics." Shorn of its fusty, turn-of-the-century diction, this statement could just as easily have come, in 2022 rather than 1897, from the mouth of a legal tech entrepreneur. To be sure, there's an element of puffery in such claims, both then and now. But even skeptics must agree that substantial change is afoot. The next few decades of the civil justice system will look very different from the past few, and many of the decisions that will determine the shape of that system are being made now. Judges, lawyers, court administrators, and policy makers should take note and ensure that the civil justice system's next, digital iteration is a healthy one. This volume seeks to equip them with the motivation and the knowledge base to do so.

[42] For Professor Rostain's prior work on these topics, see Rebecca A. Johnson & Tanina Rostain, *Tool for Surveillance or Spotlight on Inequality? Big Data and the Law*, 16 Ann. Rev. L. & Soc. Sci. 453 (2020); Tanina Rostain, *Techno-Optimism and Access to the Legal System*, 148 Daedalus 93 (2019).